

# Modeling the Context-Dependent Associations between the Gut Microbiome, Its Environment, and Host Health

Thomas J. Sharpton,<sup>a</sup> Christopher A. Gaulke<sup>b</sup>

Department of Microbiology and Department of Statistics, Oregon State University, Corvallis, Oregon, USA<sup>a</sup>; Department of Microbiology, Oregon State University, Corvallis, Oregon, USA<sup>b</sup>

**ABSTRACT** Changes in the gut microbiome are often associated with disease. One of the major goals in microbiome research is determining which components of this complex system are responsible for the observed differences in health state. Most studies apply a reductionist approach, wherein individual organisms are evaluated independently of the surrounding context of the microbiome. While such methods have yielded valuable insights into the microbiome, they fail to identify patterns that may be obscured by contextual variation. A recent report by Schubert et al. [A. M. Schubert, H. Sinani, and P. D. Schloss, *mBio* 6(4): e00974-15, 2015, doi: 10.1128/mBio.00974-15] communicates an alternative approach to the study of the microbiome's association with host health. By coupling a multifactorial experimental design with regression modeling, the authors are able to profile context-dependent changes in the microbiome and predict health status. This work underscores the value of incorporating model-based procedures into the investigation of the microbiome and illustrates the potential clinical transformations that may arise through their use.

One of the grand objectives in microbiome research is understanding how these communities of microorganisms interact with their environment. In the case of the human gut microbiome, this knowledge is clinically valuable. Changes in the host's health status or the ecology of the gut can differentially influence the taxa that comprise the microbiome and restructure the community composition. Correspondingly, changes in microbiome structure can affect the ecology of the gut or health of the host. Understanding these interactions facilitates the development of clinical diagnostics in the case that community composition or the abundances of specific taxa are predictive of disease. It can also yield novel therapeutics, such as probiotics, in the case that changes in specific taxa are shown to modulate health.

The standard approach for characterizing these interactions is to use tests of association. In this approach, the abundance of each taxonomic group is quantified in a variety of microbiome samples and then statistical tests are performed to determine whether changes in the abundance of each group are related to changes in the microbiome's environment (e.g., test of correlation). This approach simplifies the discovery of potential interactions by analyzing each group independently of the others under the assumption that meaningful interactions between taxa and their environment will be robust despite variation in the surrounding structure of the microbiome. This approach has been widely applied to microbiome studies, especially those with simple study designs (e.g., case versus control), and has yielded tremendous insight into how specific groups of taxa associate with environmental parameters, including human disease and nutritional state.

However, in complex systems, context may matter. Because the aforementioned approach isolates the analysis of each taxon from the rest, it fails to resolve complex associations that arise through the interactions among taxa that comprise the community. As a toy example, consider a taxon may that positively associate with disease when *Lactobacillus* is absent from the community but does not associate with disease when *Lactobacillus* is present. This change in association could possibly be due to the contextual nature of cellular phenotype; for example, *Lactobacil-*

*lus* may excrete a compound that changes the metabolic state of the taxon to behave in a commensal rather than pathogenic manner. The frequency of these types of interactions is not well understood, and they may explain some of the inconsistent observations made by different studies of the microbiome in similar disease models (e.g., see reference 1). If our goal is to produce clinically meaningful tools, we should seek to appropriately model these context-dependent interactions.

The discovery of robust patterns of association between microbiota and their environment is important to our ability to combat human disease, such as *Clostridium difficile* infection. *C. difficile* is an environmentally acquired enteropathogen that causes debilitating and life-threatening diarrhea. The gut microbiome appears to play an important role in protecting the host from infection. *C. difficile* rates of infection are elevated in individuals who have recently been treated with antibiotics, which has contributed to its success as a nosocomial infectious agent. Additionally, manipulation of the microbiome via transplantation from healthy donors is a highly effective form of therapy, even in cases where all other treatments have failed (2, 3). While we have learned quite a bit about the microbiome's relationship with *C. difficile* infection, there appear to be complex relationships that influence health outcomes. For example, it is not understood why antibiotic administration results in disease in some patients but not others. Nor is it understood why different types of antibiotics appear to be associated with different disease risks. It has also been difficult to determine which microbiota are associated with protecting against *C. difficile* infection. For example, antibiotic intervention is associated with a loss in *C. difficile* colonization resistance and a

Published 8 September 2015

**Citation** Sharpton TJ, Gaulke CA. 2015. Modeling the context-dependent associations between the gut microbiome, its environment, and host health. *mBio* 6(5):e01367-15. doi:10.1128/mBio.01367-15.

**Copyright** © 2015 Sharpton and Gaulke. This is an open-access article distributed under the terms of the [Creative Commons Attribution-Noncommercial-ShareAlike 3.0 Unported license](#), which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

Address correspondence to Thomas J. Sharpton, [thomas.sharpton@oregonstate.edu](mailto:thomas.sharpton@oregonstate.edu).

decrease in the abundance of *Lachnospiraceae* and *Barnesiella* (4), while microbiome transplantations are instead associated with a bloom in *Bacteroidetes* (5). These observations intimate that multiple groups of taxa may contribute to colonization resistance and that the overall context of the community may be an important factor in determining resistance.

In a recent report, Schubert et al. (6) discuss a series of investigations aimed at disentangling the complex web of interactions that comprise the microbiome in order to model the context-dependent associations between the microbiome and *Clostridium difficile* infection. The authors adopted a mouse model study design that allowed them to tune a variety of experimental parameters, including antibiotic treatment, the specific class and concentration of antibiotic, exposure to *C. difficile*, and the time between administration of the antibiotic and infection. By modulating these various parameters, the authors were able to explore how different types of community perturbations or different microbiome starting states affected the association between the abundance of specific microbiota and resistance to *C. difficile* intestinal colonization. To quantify these associations, the authors applied a machine learning technique, known as a random forest regression, to mathematically model how antibiotic exposure, structural variation in the microbiome, and time to infection interact to influence *C. difficile* colonization resistance. Briefly, the authors constructed a series of regression trees that predicted *C. difficile* colonization levels based on an assemblage of bacterial populations and experimental parameters (e.g., antibiotic class). In effect, this model uses a complex series of “if, then” statements to estimate *C. difficile* colonization levels given knowledge of the initial community composition and experimental parameters. The model constructed by the authors was able to explain ~77% of the variation observed in *C. difficile* colonization levels across samples. However, when the authors instead predicted whether an individual would be colonized by *C. difficile* or not, they observed an error rate of only 10.7%.

The analysis advanced by Schubert et al. (6) underscores the importance of considering contextual interdependencies when analyzing microbiome data. For example, they found that exposure to different classes of antibiotics resulted in different community compositions and that, within these resulting communities, *C. difficile* colonization resistance was associated with different taxa. Furthermore, they identified cases where the same taxon produces opposite correlations with *C. difficile* colonization resistance depending upon the antibiotic exposure. For example, an operational taxonomic unit (OTU) from *Bacteroides* was positively associated with *C. difficile* colonization in streptomycin-treated mice but was negatively associated in mice treated with cefoperazone. Furthermore, according to their model, considering the interactions between taxa is important for predicting colonization resistance. The authors found that decreases in OTUs associated with *Porphyromonadaceae*, *Lachnospiraceae*, *Lactobacillus*, *Alistipes*, and *Turicibacter* are associated with increased susceptibility to infection when coupled with increases in *Escherichia* or *Streptococcus* OTUs. When analyzed in isolation, many of these taxa are not predictive of disease (e.g., an OTU within *Akkermansia*) or are highly predictive under some conditions (e.g., an *Escherichia* OTU in ampicillin-treated mice) but not others (e.g., the same *Escherichia* OTU in streptomycin-treated mice). In short, Schubert et al.’s (6) regression models provide a more robust prediction of *C. difficile* colonization resistance than traditional anal-

yses would provide. Additionally, their findings indicate that the phenotype of colonization resistance is complex and depends upon a variety of parameters.

Schubert et al.’s (6) work holds important long-term implications. First, it can help advance personalized microbiome-based medicine. For example, extensions of their models may ultimately be used to screen clinical patients’ microbiomes prior to administration of antibiotics and predict their likelihood of developing an infection. Similarly, they may help doctors select patient-appropriate antibiotics or those that minimize the risk of infection. Furthermore, these models may help identify ideal microbiome donor matches to improve the efficacy of microbiome transplantations for treating *C. difficile* infection. Second, these models provide insight into those taxa that are consistently associated with *C. difficile* resistance, which can improve the development of effective probiotics based on a complex of organisms. Finally, these models provide insight into how the microbiome operates and can be used to establish hypotheses about the ecological interactions between specific groups of taxa or how perturbations to a community affect the functional relationship between taxa. Of course, the application of these models would benefit from additional investigation, such as the consideration of additional variables that may be relevant to colonization, including the initial immunological state of the host. Furthermore, it may be valuable to compare alternative regression or modeling approaches, with the aim of improving the accuracy of the predictions. Finally, the extension of these models to clinical settings requires validation in human populations, which are subject to sources of variation beyond those in an experimental mouse model, including genetic and nutritional variance.

Other studies have applied similar analytical methods to investigate the complex of dependencies that influence the microbiome (7). For example, Belzer et al. identified groups of enteric microbiota that exhibit consistent temporal patterns of variation in response to *Citrobacter rodentium*-induced colitis (10). Faith et al. applied linear models to understand how dietary variation perturbs a defined gut microbiome (8). Statnikov et al. used multivariate models to predict psoriasis using skin microbiome community profiles (9). These mathematical model-based frameworks provide insight into the mechanisms through which microbiomes operate and will likely be essential in developing a comprehensive view of how hosts and their microbiomes interact. Future investigations should similarly consider establishing study designs that enable robust, predictive modeling of the microbiome.

## REFERENCES

1. Finucane MM, Sharpston TJ, Laurent TJ, Pollard KS. 2014. A taxonomic signature of obesity in the microbiome? Getting to the guts of the matter. *PLoS One* 9:e84689. <http://dx.doi.org/10.1371/journal.pone.0084689>.
2. Kelly CP. 2012. Current strategies for management of initial *Clostridium difficile* infection. *J Hosp Med* 7(Suppl 3):S5–S10. <http://dx.doi.org/10.1002/jhm.1909>.
3. Van Nood E, Vrieze A, Nieuwdorp M, Fuentes S, Zoetendal EG, de Vos WM, Visser CE, Kuijper EJ, Bartelsman JFW, Tijssen JGP, Speelman P, Dijkgraaf GW, Keller JJ. 2013. Duodenal infusion of donor feces for recurrent *Clostridium difficile*. *N Engl J Med* 368:407–415. <http://dx.doi.org/10.1056/NEJMoa1205037>.
4. Chang JY, Antonopoulos DA, Kalra A, Tonelli A, Khalife WT, Schmidt TM, Young VB. 2008. Decreased diversity of the fecal microbiome in recurrent *Clostridium difficile*-associated diarrhea. *J Infect Dis* 197:435–438. <http://dx.doi.org/10.1086/525047>.

5. Seekatz AM, Aas J, Gessert CE, Rubin TA, Saman DM, Bakken JS, Young VB. 2014. Recovery of the gut microbiome following fecal microbiota transplantation. *mBio* 5(3):e00893-14. <http://dx.doi.org/10.1128/mBio.00893-14>.
6. Schubert AM, Sinani H, Schloss PD. 2015. Antibiotic-induced alterations of the murine gut microbiota and subsequent effects on colonization resistance against *Clostridium difficile*. *mBio* 6(4):e00974-15. <http://dx.doi.org/10.1128/mBio.00974-15>.
7. Bucci V, Xavier JB. 2014. Towards predictive models of the human gut microbiome. *J Mol Biol* 426:3907–3916. <http://dx.doi.org/10.1016/j.jmb.2014.03.017>.
8. Faith JJ, McNulty NP, Rey FE, Gordon JI. 2011. Predicting a human gut microbiota's response to diet in gnotobiotic mice. *Science* 333:101–104. <http://dx.doi.org/10.1126/science.1206025>.
9. Statnikov A, Alekseyenko AV, Li Z, Henaff M, Perez-Perez GI, Blaser MJ, Aliferis CF. 2013. Microbiomic signatures of psoriasis: feasibility and methodology comparison. *Sci Rep* 3:2620. <http://dx.doi.org/10.1038/srep02620>.
10. Belzer C, Gerber GK, Roeselers G, Delaney M, DuBois A, Liu Q, Belavusava V, Yeliseyev V, Houseman A, Onderdonk A, Cavanaugh C, Bry L. 2014. Dynamics of the microbiota in response to host infection. *PLoS One* 9:e95534. <http://dx.doi.org/10.1371/journal.pone.0095534>.

---

*The views expressed in this Commentary do not necessarily reflect the views of this journal or of ASM.*